Faculty Working Papers

CONVEX OPTIMIZATION AND LAGRANGE MULTIPLIERS

Charles E. Blair

#407

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

June 8, 1977

CONVEX OPTIMIZATION AND LAGRANGE MULTIPLIERS

Charles E. Blair

#407

Convex Optimization and Lagrange Multipliers

by

Charles E. Blair

Dept. of Business Administration

June 8, 1977

This work was supported by a grant from Investors in Business Education, University of Illinois.

Abstract

We show how the duality theorem of linear programming can be used to prove several results on general convex optimization.

Let f, $g_1,\ldots,g_k$ be convex functions defined on a convex subset S of a vector space. Let $T = \{x\in S \mid g_i(x)\leq 0 \ 1\leq i\leq k\}$. We assume throughout that T is non-empty.

We use linear programming theory to explore the relationship between the problem of minimizing $f(x)$ $x\in T$ and the Lagrange dual problem of minimizing $f(x) + \Sigma\lambda_i g_i(x)$ $x\in S$ for suitable $\lambda_i\geq 0$. This is motivated by the work of Duffin [1, 2, 3]. The main tool we shall need is a version of the duality theorem of linear programming [5, theorems 1.1.9 and 1.7.13].

Lemma: Let $A = \{x \mid Bx\geq b \text{ and } Cx=c\}$. If A is empty, there are $U\geq 0$ and V such that $U^tB+V^tC=\vec{0}$ and $Ub+Vc>0$. If A is non-empty and every $x\in A$ satisfies $dx\geq e$, then there are $U\geq 0$ and V such that $U^tB+V^tC=d$ and $Ub+Vc\geq e$.

We will not require separating hyperplane theorems or results from semi-infinite programming.

Theorem 1: $f(x)\geq L$ for every $x\in T$ if and only if for every finite $F\subset S$ there are $\lambda_1,\ldots,\lambda_k$ such that $f(x)+\Sigma\lambda_i g_i(x)\geq L$ for every $x\in F$.

Proof: The "if" part is immediate. For the "only if" part it suffices to prove the result for those finite F which contain members of T. Let $F=\{y_1,\ldots,y_N\}$ $y_1\in T$. For such F the system of equations and inequalities in unknowns $\theta_1,\ldots,\theta_N$

$$\text{(D)}\qquad \sum_{1}^{N} \theta_i = 1$$

$$\sum_{i=1}^{N} \theta_i g_j(y_i)\leq 0 \qquad 1\leq j\leq k$$

$$\theta_i \geq 0$$

has the solution $\theta_1=1$. By convexity, if θ_i is a solution to (D), $\Sigma\theta_i y_i\in T$.

Since we are assuming $f(x)\geq L$ for $x\in T$, every solution to (D) must satisfy $\Sigma\theta_i f(y_i)\geq L$. By the lemma there are $\lambda_i\geq 0$ and γ such that

$$\gamma + \sum_{j=1}^{k} \lambda_j(-g_j(y_i)) \leq f(y_i) \qquad 1\leq i\leq N$$

$$\gamma \geq L$$

so
$$f(y_i) + \sum_{j=1}^{k} \lambda_j g_j(y_i)\geq L \text{ for every } y_i\in F. \quad \text{Q.E.D.}$$

Theorem 1 may be used to prove many of the standard results on convex optimization. As an example we prove the Kuhn-Tucker theorem.

Corollary: Suppose $f(x)\geq L$ for $x\in T$ and that there is a y for which $g_i(y)<0$ $1\leq i\leq k$. Then there are $\lambda_1,\ldots,\lambda_k$ such that $f(x)+\Sigma\lambda_i g_i(x)\geq L$ for $x\in S$.

Proof: Let $\delta = \max\{g_i(y)\}$. For $x\in S$ let $A_x=\{(\lambda_1,\ldots,\lambda_k)\,|\,\lambda_i\geq 0,$ $f(x)+\Sigma\lambda_i g_i(x)\geq L$, and $-\delta(\Sigma\lambda_i)\leq f(y)-L\}$. For each $x\in S$, A_x is compact. If $H\subset S$ is finite we may use theorem 1 with $F=H\cup\{y\}$ to show $\bigcap_{x\in H} A_x$ is non-empty. Therefore, $\bigcap_{x\in S} A_x$ is non-empty, so suitable Lagrange multipliers exist. Q.E.D.

Arguments of this kind can also be used to give information about when "duality gaps" occur.

Theorem 2. There are no λ_i such that $f(x)+\Sigma\lambda_i g_i(x)\geq L$ for every $x\in S$ if and only if, for every $N>0$, there is an $x\in S$ such that $f(x)<L - N(\max_{1\leq i\leq k} g_i(x))$.

Proof: Take any $\lambda_i\geq 0$ and suppose x exists with the desired property for $N=\Sigma\lambda_i$. Then $f(x)+\Sigma\lambda_i g_i(x)\leq f(x) + (\max g_i(x))\,(\Sigma\lambda_i)<L$. So no suitable λ_i exist.

Conversely, suppose there are no suitable λ_i. For any $N>0$ and any $F=\{y_1,\ldots,y_M\}\subset S$ consider the linear system in unknowns $\lambda_1,\ldots,\lambda_k$:

$$f(y_i) + \sum_1^k \lambda_j g_j(y_i) \geq L \qquad y_i \in F$$

$$(F;N) \qquad \Sigma\lambda_i \leq N$$

$$\lambda_i \geq 0$$

If, for some N, (F;N) had a solution for every finite FCS, a compactness argument similar to that in the corollary to Theorem 1 would yield suitable multipliers λ_i. Since we are assuming such λ_i do not exist, it must be that for every N>0 there is an F such that (F;N) has no solution. By the lemma, if (F;N) has no solution, there are $\theta_1, \ldots, \theta_M \geq 0$ and $\gamma \geq 0$ such that

$$\sum_{i=1}^{M} \theta_i g_j(y_i) - \gamma \leq 0, \qquad 1 \leq j \leq k$$

and $\Sigma\theta_i(L-f(y_i)) + \gamma(-N) > 0$. By scaling, we may assume $\sum_1^M \theta_i = 1$, so that

$$\sum_{i=1}^{M} \theta_i g_j(y_i) \leq \gamma < \frac{1}{N}(L - \Sigma\theta_i f(y_i)).$$

If we take $x = \Sigma\theta_i y_i$ $\gamma \geq g_j(x)$, $1 \leq j \leq k$ and $f(x) < L - N\gamma$ follow by convexity of f and g_j. Q.E.D.

Corollary: (Compare [1], cor. 5; [2], thm. 3): Let $h(\epsilon) = \inf\{f(x) \mid g_j(x) \leq \epsilon,$ $1 \leq j \leq k$. If there are $\delta > 0$, L such that $h(x) > L$ for $0 \leq x \leq \delta$, then there are $\lambda_1, \ldots, \lambda_k$ such that $f(x) + \Sigma\lambda_i g_i(x) \geq L$ $x \in S$.

Proof: If there is an x for which $g_j(x) < 0$ the existence of suitable λ_i follows from the Kuhn-Tucker theorem,* so we assume this is not the case. h is a convex monotone function which, on our assumptions, is defined only for non-negative arguments. For $\epsilon \geq \delta$, $h(\epsilon) - L > h(\epsilon) - h(0) \geq \epsilon(\frac{1}{\delta})(h(\delta) - h(0))$. Hence for $x \in S$, $f(x) - L \geq (\max\, g_i(x))(\min\, 0, \frac{1}{\delta}(h(\delta) - h(0)))$. (Note that our assumptions

---

*We use the Kuhn-Tucker theorem for brevity. The result could be proved from Theorem 2 alone.

imply max $g_j(x)\geq 0$.) Since the condition given by Theorem 2 fails for $N=\max(0, \frac{1}{\delta}(h(0)-h(\delta))$, suitable λ_i exist.) Q.E.D.

Finally, we use a variation of these techniques to strengthen a recent result of Duffin and Jeroslow [4].

Theorem 3: Let $S=R^n$. Assume that for $\lambda_i>0$, $f(x)+\Sigma\lambda_i g_i(x)\geq L$, $x\in S$. Then there are affine functions $h_i(x)=a_i x+b_i$ ($a_i\in R^n$, $b_i\in R$) such that $h_i(x)\leq g_i(x)$ and $f(x)+\Sigma\lambda_i h_i(x)\geq L$, $x\in S$.

Proof: For $y\in S$, let $T_y=\{(h_1,\ldots,h_k)\mid h_i$ affine, $h(y)\leq g(y)$, and $f(y)+\Sigma\lambda_i h_i(y)\geq L\}$.

We identify each function $h_i(x)=a_i x+b_i$ with the ordered pair $(a_i,b_i)\in R^{n+1}$. Thus, T_y is a closed subset of $R^{(n+1)k}$. We first show that, for any finite $F\subset S$ $\bigcap_{y\in F} T_y$ is non-empty. A member of $\bigcap_{y\in F} T_y$ would be a solution to the linear inequality system in unknowns $a_1,\ldots,a_k$; $b_1,\ldots,b_k$.

$$a_i y+b_i\leq g_i(y) \qquad 1\leq i\leq k; \quad y\in F$$

$$\text{(E)} \qquad \sum_{i=1}^{k} \lambda_i(a_i y+b_i)\geq M-f(y) \qquad y\in F$$

By the lemma, (E) has no solution only if there are scalars W_{iy}, $V_y\geq 0$ such that

$$\text{(i)} \qquad \sum_{y\in F} W_{iy}y = \lambda_i \sum_{y\in F} V_y y, \quad 1\leq i\leq k$$

$$\text{(ii)} \qquad \sum_{y\in F} W_{iy} = \lambda_i \sum_{y\in F} V_y, \quad 1\leq i\leq k$$

$$\text{(iii)} \qquad \sum_{\substack{1\leq i\leq k\\ y\in F}} W_{iy}g_i(y) < \sum_{y\in F} V_y(M-f(y)).$$

If there were W,V satisfying (i)-(iii) we could set $V_y'=V_y/\Sigma V_y$ and $W_{iy}'=W_{iy}/\lambda_i\Sigma V_y$ so that (i)-(iii) would be satisfied and, by (ii), $\Sigma V_y'=\sum_{y\in F} W_{iy}'=1$, $1\leq i\leq k$. Condition (i) becomes $\Sigma W_{iy}'y=\Sigma V_y'y$. Condition (iii) becomes $\sum_{y\in F} V_y'f(y)+\Sigma\lambda_i(\sum_{y\in F} W_{iy}'g_i(y))<M$. If $z=\Sigma W_{iy}'y=\Sigma V_y'y$ this implies, by convexity, $f(z)+\Sigma\lambda_i g_i(z)<M$

which contradicts our assumption about the λ_i. Therefore (E) has solutions for every finite F.

To complete the proof we must show $\bigcap_{y\in S} T_y$ is non-empty. Let e_j=jth unit vector. We show that if F contains $\pm e_j$ $1\le j\le n$ and the zero vector, then $\bigcap_{y\in F} T_y$ is bounded. Since each T_y is closed, compactness yields the desired result. For $1\le i\le k$ we must have $h_i(\vec{0})=b_i\le g_i(0)$, $f(\vec{0})+\Sigma\lambda_i b_i\ge L$, $h_i(e_j)=$(jth component of a_i) + $b_i\le g_i(e_j)$, and $h_i(-e_j)\le g_i(-e_j)$. Since all the λ_i are positive this implies bounds on a_i, b_i. Q.E.D.

Our proof of Theorem 3 works for any convex $S\subset R^n$ which includes $\pm e_j$ and $\vec{0}$. By suitable translations, this implies the results for any fully dimensional convex $S\subset R^n$. Further modifications yield the result for arbitrary convex $S\subset R^n$.

I would like to thank Richard Duffin and Robert Jeroslow for their encouragement, and for supplying me "sneak previews" of [2] and [4].

REFERENCES

1. Duffin, R.J. "Convex Analysis Treated by Linear Programming." Mathematical Programming, 4, pp. 125-43.

2. ________. "The Lagrange Multiplier Method for Convex Programming." Proceedings of National Academy of Sciences, 72, pp. 1778-1781.

3. ________. "Convex Programming Having Some Linear Constraints." Proceedings of National Academy of Sciences, 74, pp. 26-8.

4. Duffin, R.J. and Jeroslow, R.G. Private communication.

5. Stoer, J. and Witzgall, C. Convexity and Optimization in Finite Dimensions I. Springer-Verlag, 1970.